Cloning and Sequence Analysis of the *Escherichia coli* metH Gene Encoding Cobalamin-dependent Methionine Synthase and Isolation of a Tryptic Fragment Containing the Cobalamin-binding Domain*

(Rceived for publication, October 21, 1988, and in revised form, March 22, 1989)

Ruma V. Banerjee, Nancy L. Johnston, James Kenneth Sobeski, Prasanta Datta, and Rowena G. Matthews‡

From the Biophysics Research Division and the Department of Biological Chemistry, the University of Michigan, Ann Arbor, Michigan 48109

A gene encoding cobalamin-dependent methionine synthase (EC 2.1.1.13) has been isolated from a plasmid library of *Escherichia coli* K-12 DNA by complementation to methionine prototrophy in an *E. coli* strain lacking both cobalamin-dependent and -independent methionine synthase activities (RK4536: metE, metH). Maxicell expression of a series of plasmids containing deletions in the metH structural gene was employed to map the position and orientation of the gene on the cloned DNA fragment. A 6.3-kilobase EcoRI-SalI fragment containing the gene was cloned into the sequencing vector pGEM3B for double-stranded DNA sequencing; the MetH coding region consists of 5372 nucleotides. The enzyme was purified from an overproducing strain of *E. coli* harboring the recombinant plasmid, in which the level of methionine synthase was elevated 30- to 40-fold over wild-type *E. coli*. Recombinant enzyme is a protein of 123,640 molecular weight and has a turnover number of 1,450 min⁻¹ in the standard assay. These values are to be compared with previously reported values of 133,000 for the molecular weight and 1,240–1,560 min⁻¹ for the turnover number of the homogenous enzyme purified from a wild-type strain of *E. coli* B (Frasca, V., Banerjee, R. V., Dunham, W. R., Sands, R. H., and Matthews, R. G. (1988) *Biochemistry* 27, 8458–8465). Limited proteolysis of the native enzyme with trypsin resulted in loss of enzyme activity but retention of bound cobalamin on a peak of 28,000 molecular weight. This fragment has been shown to extend from residue 643 to residue 900 of the 1124-residue deduced amino acid sequence.

The terminal step in the *de novo* biosynthesis of methionine in *Escherichia coli* involves a transmethylation of CH₃H₄folate to homocysteine. The reaction is catalyzed by two forms of methionine synthase: the cobalamin-dependent (MetH) and the cobalamin-independent (MetE) enzymes. The methyl donor for both enzymes is CH₃H₄folate, produced by the metE gene product at a point of convergence of two major pathways, the methionine biosynthetic pathway and the folate or C₁ pathway. While the metE gene product requires the triglutamate form of the folate substrate, the metH gene product can utilize either the mono- or triglutamate forms. During aerobic growth, formation of the active MetH holoenzyme requires an exogenous supply of cobalamin as *E. coli* are unable to synthesize the prosthetic group aerobically. Our laboratory has been engaged in studying the mechanism of the reaction catalyzed by the cobalamin-dependent methionine synthase. However, the availability of only small amounts of pure enzyme has been a limiting factor. Thus, as a first step to ensure greater availability of the enzyme, cloning, sequencing, and expression of the cobalamin-dependent methionine synthase (EC 2.1.1.13) from *E. coli* K-12 were undertaken.

We report here the cloning of the metH gene by screening a plasmid library of *E. coli* DNA by complementation of strain RK4536 (metE−, metH+), which requires methionine for growth, to Met⁺ in the presence of hydroxocobalamin. From the nucleotide sequence data, together with enzyme activity assays and maxicell expression, the cloned complementing activity was positively identified as the structural metH gene. Limited proteolysis of methionine synthase was attempted in an effort to delineate the region responsible for cobalamin binding in the primary sequence. The orange-pink color associated initially with the intact protein was found to be associated with a fragment of 28,000 molecular weight after tryptic digestion of the native enzyme. N-terminal sequences of tryptic fragments confirm portions of the deduced amino acid sequence and establish the boundaries of the cobalamin-binding region.

**EXPERIMENTAL PROCEDURES**

**Materials**—Tryptone and yeast extract were supplied by Difco. Agarose, lysozyme, T4 DNA ligase, ribonuclease, and all restriction enzymes were obtained from Bethesda Research Laboratories. Low melting Sea Plaque agarose was from FMC Bioproducts. The Gene Clean Kit from Bio 101 was employed to recover DNA from agarose gels. The Erase-A-Base and the K/RT sequencing kits were purchased from Promega. [⁷⁵S]Methionine (900 Ci/mmol) and EN3HANCE were from Du Pont-New England Nuclear Research Products and d-[⁷⁵S]dATP (400 Ci/mmol) was from Amersham. The following items were bought from Sigma: L-α-amino acids, nucleotides, hydroxy-β-cyano, and methylcobalamin, sodium ampicillin, β-cyclodextrin, dithiothreitol, phenylmethylsulfonyl fluoride, TLCK, EDTA, DEAE-Sepharose (fast flow), trypsin (treated with L-1-toyamido-2-phenylethyl chloromethyl ketone), and ethidium bromide.

13888
**Cobalamin-dependent Methionine Synthase**

**Bacterial Strains and Plasmids—**All bacterial strains used are derivatives of E. coli K-12 and are reported in Table I. Plasmid pUC8 has been described previously (Messing and Viera, 1982). Plasmids pGEM3B and pGEM4B were from Promega.

**Media—**LB broth, LB agar, and M9 minimal media were prepared as described by Maniatis et al. (1982). Minimal MOPS was prepared as described by Neidhardt et al. (1974). MOPS minimal medium was supplemented with 0.4% glucose and thiamine. The concentrations of amino acid and purine supplements were those used in the defined rich medium described by Wanner et al. (1977). Other supplements were added at the following concentrations: ampicillin, 100 μg/ml in liquid cultures; 10 μg/ml in plated culture, cobalamin, 1 μM; and thiamine, 10 μM. Supplements to glucose minimal media for growth of RK4536 were leucine, proline, tryptophan, lysine, arginine, methionine, adenosine, and guanine.

**DNA Preparation—**Plasmid DNA was prepared on a small scale as described in the Promega K/RT sequencing technical manual, while large scale purifications were done as described by Maniatis et al. (1982). DNA restriction fragments were isolated from 1% low melting agarose gels and purified by adsorption to “glass milk” as specified by the supplier of the Gene Clean kit.

**MaxiX Expression—**The maxiX strain CSR603 was transformed with nested deletions constructed from pBl6.3 and containing variable lengths of the metH-containing insert, and transformants were selected by their resistance to ampicillin. Expression of plasmid-encoded proteins in maxiX cells was accomplished as described by Silhavy et al. (1984), and the [35S]methionine-labeled proteins were analyzed on either silver or Coomassie blue gels (1984) according to the protocol of Blumenthal et al. (1976) or on 12% polyacrylamide gels containing SDS. The gels were fixed overnight in a solution containing 25% isopropanol alcohol, 10% acetic acid, and 1% trichloroacetic acid. Fluorography was done by immersing the gels in ENHANCE with shaking for 30 min followed by shaking in cold water for 30 min. The gels were dried under vacuum at 60 °C for 2.5 h prior to exposure.

**Generation of Nested Deletions and Nucleotide Sequence Analysis—**A series of nested deletions were generated with exonuclease III by the procedure described by Henikoff (1984) as recommended by the supplier of the Erase-A-Base kit. The digestion was done at 35 °C, and samples were removed every 30 s. The estimated rate of digestion under these conditions was approximately 400 bases/min. Double-stranded nucleotide sequence analysis was done by the dideoxy-nucleotide method of Sanger et al. (1977), using α-[35S]dATP as recommended by the manufacturer of the K/R sequencing kit. Klenow polymerase was employed for the sequencing reactions at 55 °C. Sequence of the transcribed strand was determined with the SP6 primer for pBl6.3 deletion subclones. Sequence analysis of the opposite strand was accomplished by extending synthetic primers (17mers) that had been synthesized on an ABI Model 380A DNA Synthesizer at the DNA Sequencing Facility, Dept. of Biological Chemistry, the University of Michigan.

**Enzyme Purification and Assay—**Enzyme activity was monitored in crude extracts of recombinant bacteria containing pBl6.3 as follows. Six 1-liter cultures of DH5αF′/pBl6.3 in glucose minimal medium (M9) supplemented with cobalamin (1 μM), micronutrients (containing molybdate, borate, Cu2+, Mn2+, and Zn2+ as described by Neidhardt et al., 1974), and ampicillin were grown to an absorbance at 620 nm of 3.3. The cultures were protected from light during cycles to prevent overheating. The suspension was ultracentrifuged at 100,000 × g for 1 h to remove cell debris and unbroken cells. Samples were removed for enzyme activity and protein assays. The orange-pink supernatant fluid was immediately loaded onto two DEAE-Sepharose columns (2.2 × 20 cm) equilibrated with 180 mM KPi. Each column was washed with 100 ml of 180 mM KPi, containing 1 μM adenosylmethionine, and the protein was eluted with a 500-ml linear gradient from 180 to 500 mM KPi, containing 1 μM adenosylmethionine. Fractions were monitored for enzyme activity by the assay that has been described previously (Prascha et al., 1988). Enzyme-containing fractions were concentrated and dialyzed overnight against 25 mM KPi, and purified by FPLC on a Mono Q HR (16/10, from Pharmacia LKB Biotechnology) column as described previously (Frasca et al., 1988). Fractions were inspected visually, and those with the pink color of methionine synthase were pooled and concentrated. One unit of methionine synthase activity catalyzed the formation of 1 pmol of methionine/min at 37 °C. The turnover number was estimated as micromoles of methionine formed (min)-1 (μmol of cobalamin)-1. The cobalamin content of the enzyme was estimated spectrophotometrically using an ε495 of 11,000 μM⁻¹.

**Separation of Tryptic Peptides—**The specific conditions employed for the trypsin digests were described in the legends of Figs. 6 and 7. Tryptic fragments of methionine synthase were isolated on one tube of the two FPLC separation columns in a Sepharose Fast Flow column (from Pharmacia). The protein was eluted isocratically with 50 mM KPi, pH 7.2, containing 500 mM KCl, and the protein fractions were concentrated and desalted in Centricon 30 microconcentrators. Alternatively, the fragments after proteolysis were separated on native gels (lacking detergents) by electrophoresis at low voltage (100 V) so as to prevent the buffer from overheating. Electrophoretic (in an apparatus from Schleicher and Schuell) of the bands from native gels was performed overnight at 100 V in TAE buffer (40 mM Tris acetate, pH 8.2, 1 mM EDTA). The protein was then desalted on an EconoPac 10DG column from Bio-Rad and concentrated by lyophilization in a Speed Vac.

**N-terminal Sequence Determination—**The N-terminal sequences of the enzyme purified from E. coli B and of the tryptic peptides from the recombinant enzyme were determined by gas phase sequence analysis at the UM Protein Sequencing Facility, University of Michigan, Ann Arbor, MI.

**Computational Methods—**Compilation of sequence overlap was accomplished with the DM sequence analysis program written by David W. Mount and Bruce Conrad, University of Arizona, Tucson (Mount and Conrad, 1986). Codon translation and open reading frames were plotted with the same program. Codon preference and test analyses based on the alignments were performed by Fickett (1984) and Fickett, respectively, were plotted with programs in the Wisconsin GCG package (Devereaux et al., 1984). The pl of the deduced amino acid sequence was calculated by using an algorithm developed by Alex Pertsemidis, Department of Biological Chemistry, the University of Michigan. Quantitative estimation of protein bands on analytical gels was accomplished by digital conversion of data with a video acquisition system with a CCD camera and analyzed with an interactive program written by Forrestor by Shawn Williams (Biophysics Research Division, the University of Michigan) for a Silicon Graphics Iris 2900 T computer.

**RESULTS**

**Cloning of the methH Gene—**To construct a plasmid library of E. coli genome, random fragments of 5 to 10 kb pairs of

---

**Table I**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>RK4536</td>
<td>leuB6, proC32, trpE38, lysA23, argH1 metE70, metH156, purE42, aicR2, tonA23, lacZ56, terX7, tonA23, gyrA90, rpsL109, yf-5, χ-mtl-1, rpoB308, thi-1</td>
<td>B. Bachmann, Yale University</td>
</tr>
<tr>
<td>DH5αF′</td>
<td>ϕ80daciZAM15, endA1, recA1, gyrA, hadR17 (α, ma2), chromosomal thi-1, χ-relA</td>
<td>Bethesda Research Laboratories</td>
</tr>
<tr>
<td>CSR603</td>
<td>waaA6, recA1, rpsL, X′, thr, leu pro, arg, thi-1</td>
<td>F. C. Neidhardt, University of Michigan</td>
</tr>
</tbody>
</table>
DNA, isolated by partial Sau3A digestion followed by sucrose density gradient centrifugation, were cloned into the BamHI site of the plasmid pUC8. The metH gene was cloned by transforming *E. coli* strain RK4536 (metE70, metH156) with the plasmid library and selecting for Ap'Met⁺ transformants on minimal medium plates containing hydroxocobalamin and other supplements but lacking methionine. Of the nine transformants found, seven showed a cobalamin requirement and harbored the parental markers. The plasmid (designated pMH4) from one such transformant was isolated by the alkali method and purified on a CsCl-ethidium bromide gradient. Upon retransformation of RK4536 with purified DNA, a large number of Ap'Met⁺ colonies appeared indicating probable cloning of the metH gene.

The plasmid pMH4 contained an 11.5-kb insert with two EcoRI sites and a SalI site (Fig. 1). The 6.3-kb EcoRI-SalI fragment was isolated and cloned in opposite orientations into the vectors pGEM3B and pGEM4B. The resulting constructs, named p3B6.3 and p4B6.3, respectively, complemented RK4536 to methionine independence, suggesting the presence of the entire metH coding region. A further attempt to subclone the 5.2-kb EcoRI-HindIII fragment was not successful.

**Identification of the metH Gene Product**—In order to ascertain whether p3B6.3 indeed contained the intact structural gene, maxicell expression of the plasmid-encoded proteins was carried out. The labeled proteins were separated on two-dimensional gels (data not shown). Two major spots were seen corresponding to the vector-encoded β-lactamase and the insert-encoded MetH. The latter co-migrated with methionine synthase from wild type *E. coli* (B or K-12) cells that have been previously indexed at C137 (migrating at 95 × 120) in the gene-protein index (Frasca et al., 1988).

**Location of the metH Gene within p3B6.3**—Methionine synthase (MetH) isolated from *E. coli* B has an estimated monomeric subunit size of 133,000 daltons (Frasca et al., 1988). Hence, the estimated size of the coding region is ~3.8 kb. In order to estimate the approximate location of the gene in the 6.3-kb insert of p3B6.3 and also as a sequence analysis strategy, a series of nested deletions were generated. The unique SalI and SphI sites were employed to open the plasmid with ends appropriate for unidirectional exonuclease III digestion. It is important to note that cleavage at the PstI (which like SphI generates 3' overhangs) led to failure to cleave at the SalI site, presumably because the two sites are contiguous in the polylinker region.

The deletion subclones were used to transform DH5αF'. Plasmid DNA from a few selected nested deletions, containing variable insert lengths, was used to transform the maxicell strain CSR603. Maxicell expression of these transformants revealed a series of truncated proteins (Fig. 2). Since the deletions had been generated in a counterclockwise direction, from the SalI end, this result unambiguously established the clockwise direction of transcription of the metH gene. Truncated proteins were produced because the flanking vector sequence contained termination codons in all three reading frames within the first 45 bases, with the stop codons lying 15, 24, and 45 bases downstream of the insert. Hence, the maximum error due to readthrough of vector sequence could be ~1650 daltons. By comparing the sizes of the truncated polypeptides to the sizes of the shortened inserts, the approximate site of translation initiation of the metH gene could be predicted. It is not known why the 5.3-kb insert failed to show methionine synthase expression, whereas expression was seen with inserts both immediately larger and smaller.

**Nucleotide Sequence of metH**—The strategy employed for nucleotide sequence analysis is shown in Fig. 3. Double-stranded nucleotide sequence analysis with Klenow polymerase was accomplished by the dideoxy chain-terminating method. The sequence of the transcribed strand was determined by using the SP6 (for p3B6.3 subclones) primer. The subclones were staggered by ~200 bases. The sequence of the opposite strand was determined by primer extension, with primers (17-mers) complementary to the transcribed strand having been synthesized at ~180-base pair intervals. The

---

**Fig. 1.** Restriction map of pMH4 and construction of plasmids p3B6.3 and p4B6.3. Details of the construction are given under “Experimental Procedures.” The EcoRI and SalI sites used for subcloning are underlined.

**Fig. 2.** Maxicell expression of full length and truncated metH gene products from a series of deletion subclones of p3B6.3. Details are described under “Experimental Procedures.”
sequences of both strands were completely determined including 220 bases upstream and 180 bases downstream of the protein coding sequence. Fig. 4 shows the complete nucleotide sequence of the metH gene. A 3372-base open reading frame starts with a GTG initiating codon and follows a ribosomal binding site GGAG, which fits the consensus Shine-Dalgarno sequence (Gold and Stormo, 1987). The first 11 residues of the N-terminal amino acid sequence of the protein from E. coli B were identical with the deduced amino acid sequence as shown in Fig. 4. The initiator methionine is presumably removed from the mature peptide. The open reading frame encodes 1124 amino acids, with a predicted molecular mass of 123,640 daltons, within 9% of the value given by SDS-polyacrylamide gel electrophoresis. The predicted amino acid composition and the experimental value obtained from total amino acid hydrolysis of the homogenous protein from E. coli B are in good agreement (data not shown). The codon usage probability (data not shown) is high in only one of the three frames and supports the verity of the sequence determination. It yields a codon preference statistic of 0.929, a value similar to those for other biosynthetic genes. This overall codon preference statistic of metH puts it in a category between (viz. lacI) and moderately expressed (viz. rpoB) genes. The test code analysis (data not shown) which distinguishes coding from noncoding sequences and is independent of reading frame, and therefore frame shift-independent, similarly makes a prediction for a protein coding sequence with boundaries corresponding to that deduced from the sequence translation. The pl of the deduced amino acid sequence was found to be 5.3 ± 0.2, in close agreement with a pl of 5.2 obtained from denaturing O'Farrell gels, again pointing to the accuracy of the DNA sequence determined. Random DNA sequences generate pl values that are much more basic. Comparison of the entire metH sequence with both DNA and protein sequence available in the GenBank and NBRF databases and more specifically with folate- and cobalamin-binding proteins did not yield any significant homologies.

Enzyme Purification—DH5αF′ was transformed with plasmid pB6.3, and the transformants were selected for their Ap′ phenotype on rich media. A 6-liter minimal medium culture supplemented with cobalamin (methyl- or cyano-), ampicillin, and micronutrients yielded cells that were pale pink in color. The orange-pink color of the suspension after cell wall disruption was due mostly to protein-bound cobalamin, and the extract had a specific activity of ~0.4 μmol (min)⁻¹ (mg of protein)⁻¹. Quantitative estimation of the methionine synthase band separated from the extract on an SDS-polyacrylamide gel showed that it comprised ~5 to 7% of the total cellular protein. The enzyme was overproduced 30- to 40-fold based on the specific activity in the extract with respect to wild type E. coli B cells, presumably by virtue of its gene being on a multicopy plasmid. The high yield of the enzyme in a strain harboring the cloned metH gene facilitated large scale enzyme isolation. Chromatography on a DEAE-Sepharose column yielded protein that was 80 to 90% pure (Fig. 5) and was obtained in >90% yield. This anion exchange resin has a much higher flow rate than the previously used DEAE-Sephadex resin (Frascia et al., 1988) and permitted a rapid purification of the enzyme. The specific activity of the enzyme at this stage was 4.4–5.8 μmol (min)⁻¹ (mg of protein)⁻¹. Following FPLC purification, the specific activity increased to 6.6–7.8 μmol (min)⁻¹ (mg of protein)⁻¹ corresponding to a turnover number of ~1450 μmol (min)⁻¹ (μmol of enzyme-bound cobalamin)⁻¹. The purified protein was judged to be homogenous by gel electrophoresis and was obtained in an overall yield of ~60%. Approximately 6 mg of methionine synthase was isolated per liter of culture. The cobalamin derivative used to supplement the growth medium did not appear to influence the spectral properties of the isolated enzyme, which was a mixture of aquo-, methyl-, and cob(II)alamin. It could be readily converted to the cob(II)alamin form by treatment with homocysteine and dithiothreitol which were subsequently removed by filtration in Centricron 30 microconcentrators.

Isolation of Cobalamin-binding Fragment—Cleavage of methionine synthase by trypsin occurs initially at a hypersensitive site resulting in two fragments with apparent molecular masses of 95 and 35 kDa, respectively, and a corresponding loss of enzyme activity. The two fragments obtained after the first cleavage could be isolated by FPLC. Whereas the 35-kDa fragment is relatively stable, the 95-kDa peptide is further cleaved to 68- and 28-kDa fragments. The 68-kDa fragment is fairly unstable as evidenced by its relatively low intensity on analytical gels, but a 58-kDa fragment, presumably a product, accumulates. Fig. 6 shows the kinetics of formation of the various fragments and traces the origin of the 28-kDa fragment (which harbors the cobalamin-binding site as discussed below) to the 95-kDa and not to the 35-kDa peptide. The kinetics of disappearance of the 95-kDa fragment is nicely paralleled by the appearance of the 58- and 28-kDa species, whereas the intensity of the 35-kDa fragment does not change significantly.

The fate of the cobalamin-bound tryptic peptide could be followed by the migration on a native gel of a pink band that moved more rapidly than the undigested, intact protein. As the 28-kDa fragment is not formed in large amounts, monitoring its presence on unstained native gels and subsequent recovery by electroelution were greatly facilitated by the ready availability of large amounts of protein. Methionine synthase migrates to the junction of the running and stacking gels as shown in Fig. 7A, as does the 95-kDa fragment. The behavior of the 95-kDa and 35-kDa fragments on the native gel was established by their isolation by FPLC under non-denaturing conditions followed by electrophoresis under identical conditions (data not shown). The 95-kDa fragment as isolated by FPLC retained the pink color and spectral properties of the intact enzyme and provides strong evidence for the separation and retention of the cobalamin-binding domain following the initial cleavage. The identity of the 28-kDa peptide was established by electrophoresis of the electroeluted peptide on an SDS gel. The sample after electroelution, desalting, and concentration retained its pink color indicating that the enzyme remained bound to the peptide. None of the three cobalamin derivatives (methyl-, cyano-, and hydroxo-) ex-
Cobalamin-dependent Methionine Synthase

**Fig. 4.** Sequences of the *metH* gene and the predicted protein and N-terminal sequence derived from purified methionine synthase from *E. coli*. The *metH* gene sequence is numbered in *E. coli* chromosome is believed to be 91 min (Bachmann, 1987). The putative -10 and -35 regions are underlined. A region of dyad symmetry that could form a stem loop structure immediately downstream from the stop codon is indicated. S.D., Shine-Dalgarno.
Samples were applied to a discontinuous gel natured by boiling for 2 min in a solution containing 5 mM 2-mercaptoethanol, 2% SDS, 10% glycerol, and 62.5 mM Tris-HCl, pH 6.8. Samples were applied to a discontinuous gel (4% polyacrylamide stacking gel, 12% polyacrylamide running gel). After electrophoresis, the gel was stained with Coomassie Blue, destained, and dried on a filter paper. Each lane contained 16 µg of methionine synthase. Lane A, enzyme after chromatography on a DEAE-Sephacel; lane B, enzyme after FPLC purification; lane C, same as B, but with 32 µg of methionine synthase. The numbers on the left indicate the molecular mass of the standards in kilodaltons.

N-terminal sequence analysis of the 28- and 35-kDa fragments established the limits of these peptides and provided evidence for the verity of the deduced amino acid sequence (Fig. 4). The hypersensitive site between residues 900 and 901 resulting in the formation of the 35-kDa fragment lies in one of the most hydrophilic internal regions of the protein as predicted by the Peptide Structure program with the method of Kyte and Doolittle (Kyte and Doolittle, 1982). The 28-kDa fragment with an N-terminal sequence of T-D-D (beginning at residue 643 of the deduced amino acid sequence) is situated immediately upstream of the 35-kDa fragment. It is characterized by the prevalence of hydrophobic residues. Also noteworthy are the high concentration of lysines and threonine (relative to their abundance on the complete chain) and the low cysteine content. A single cysteine is present in this fragment. The position of the cobalamin-binding domain immediately upstream of the 35-kDa fragment and the complete loss of enzyme activity following cleavage at the hypersensitive site suggest that the downstream region is also required for catalysis. Proteolysis in the presence of CH₃H₂folate does not affect the initial rate of cleavage but stabilizes the 68-kDa fragment (data not shown). Proteolysis in the presence of both homocysteine and CH₃H₂folate greatly destabilizes the 95-kDa fragment. These data are consistent with the substrates interacting with the parent 95-kDa peptide which also harbors the cobalamin-binding site.

The size of the 28-kDa fragment agrees very well with the apparent size of that fragment (29 kDa) as estimated by gel electrophoresis. On the other hand, the fragment with the apparent mass of 35 kDa has an actual mass of 26 kDa. The discrepancy in the migration of this fragment is not understood. Since the "35-kDa" fragment lies at the C terminus of the protein, the possibility of a sequence error leading to a predicted premature termination was considered. The sequence in the region of the gene containing the stop codon

amined migrated into the native gel. Hence, the identification of both a parent (95-kDa) and a daughter (28-kDa) peptide fragment that retain the pink color of the native enzyme is very strong evidence for the isolation of peptides that retain the cobalamin-binding site.
and approximately 200 nucleotides upstream of it was determined with two different clones on one strand and two primers on the opposite strand. In each case, the same sequence was read, making it quite unlikely that a sequence determination error has been made.

**DISCUSSION**

The enzyme cobalamin-dependent methionine synthase has been isolated from an overproducing strain of *E. coli* bearing p4B6.3. A nutritional auxotroph of *E. coli* was used to isolate a recombinant clone that complemented the mutation and encoded cobalamin-dependent methionine synthase (Banerjee et al., 1988). The identity of the cloned gene has been confirmed by nucleotide sequence analysis, maxicell expression, and enzyme activity assays. Cloning of the *metH* and *metE* genes has been recently reported independently by another group (Old et al., 1988). They employed Tn1000 insertion-mutagenesis to map and determine the direction of transcription of the *metH* gene. From the single and slightly truncated polypeptide identified, they correctly assumed the clockwise direction of transcription from the EcoRI site. Moreover, the approximate map positions of the HindIII sites, in addition to their inability to subclone using the HindIII sites, are in agreement with our data.

Methionine synthase is a protein of 123,640-dalton molecular mass with a calculated pI of 5.3 ± 0.2. The turnover number of the purified enzyme is 1450 min⁻¹ and is within the range of fully active enzyme isolated from *E. coli* B (Frasca et al., 1988). The specific activity of the purified recombinant enzyme is, however, only 60 to 80% of that found with *E. coli* B methionine synthase (9.3–11.7 μmol (min)⁻¹ (mg of protein)⁻¹) and suggests that a portion of the enzyme is present as apoenzyme, which is inactive under standard assay conditions. It is possible that the bacteria are unable to transport cobalamin at a rate commensurate with the elevated level of apoenzyme synthesis. The growth medium is pink even after the cells have been harvested by centrifugation, indicating that the cobalamin concentration is not limiting.

**Fig. 8.** Secondary structure comparisons of segments of three cobalamin-binding proteins. Upper panel, residues 800 to 900 of MetH; middle, residues 96 to 196 of BtuR; lower, residues 505 to 605 of MutB. CF, Chou-Fasman; GOR, Garnier-Osguthorpe-Robson. The thickness of the black bars is proportional to the strength of the prediction.
Methionine synthase is one of the largest peptides known in \textit{E. coli}. Hence, the isolation of a fragment that is less than 25\% the size of the intact protein and harbors the cobalamin-binding site will facilitate studies of the interactions of the cobalamin with the apoprotein. This is especially so in light of the fact that no cobalamin-binding proteins have been crystallized to date, and the nature of a "cobalamin pocket" therefore remains unknown. The sequence appears to contain a region of moderate hydrophobicity and is flanked by extended hydrophilic segments. This retains the tertiary structure and ability to bind cobalamin after separation from adjacent peptide segments.

A comparison of both the full length methionine synthase and of the 28-kDa peptide with other cobalamin-binding proteins BtuR (Lundrigan and Kadner, 1989), BtuB (Heller and Kadaer, 1985), and MutA and MutB (Marsh \textit{et al.}, 1989) was attempted and failed to pick up any significant homologies. BtuB is the cobalamin receptor protein in the outer membrane of \textit{E. coli}, while BtuR is a soluble protein from \textit{E. coli} that has been implicated in metabolism of adenosylcobalamin. MutA and MutB represent the two subunits of methylmalonyl-CoA mutase from \textit{Propionibacterium shermanii}. MutB shows a striking primary sequence homology with the human methylmalonyl-CoA mutase.

It is possible that, as with the heme pocket, similarities may exist instead in the secondary and tertiary structures in which functionally critical regions of the structure are conserved. Comparisons within the globin family reveal that although its members have very different amino acid sequences, they share remarkably similar secondary and tertiary structures (Lesk and Chothia, 1980). Eight helices assemble in a common pattern, enclosing the heme group in pockets of similar geometry made up from homologous portions of the molecules. Only 5 positions out of the 116 that are involved in conserved interactions have the same residue in all globins, and even they are not contiguous on the primary sequence. A similar comparison of hemoglobin and cytochrome \textit{b}$_{5}$ is remarkable not only for the close superposition of the heme groups but also in the conservation of structurally equivalent units (Rossman and Argos, 1975).

Examination of the Chou-Fasman (Chou and Fasman, 1978) predictions of secondary structure for BtuR, MutB, and the 28-kDa fragment reveals a pattern of alternating \(\alpha\) helices and \(\beta\) sheets. This is distinct from the secondary structure elements involved in heme binding, which are primarily \(\alpha\) helices. The region of greatest similarity in each of the three proteins is shown in Fig. 8. The spacing and the alternation of the \(\alpha\) helices and \(\beta\) sheets in the three proteins is quite striking and is suggestive of a Rossman fold type of structure. Dimethylbenzimidazole is an unusual nucleotide that is part of the cobalamin structure. A comparison of the benzimidazole-ribosuransyl portion of cobalamin with the adenine ribose of NAD (bound to crystalline malate dehydrogenase) revealed a good alignment of the 2 molecules with a calculated root-mean-square distance between corresponding atoms of 1.05 Å (Dieckgraefe \textit{et al.}, 1988). The region of identified secondary structure similarity in the cobalamin binding proteins, shown in Fig. 8, may be involved in docking the dimethylbenzimidazolyl group.

The isolation and subcloning of the \textit{meth} gene allow for the convenient purification of methionine synthase in large amounts. This sets the stage for further characterization of the structural and mechanistic properties of the enzyme.

\textbf{Acknowledgments}—We gratefully acknowledge the generous advice and helpful suggestions from Dr. Martha Ludwig, Dr. Herbert Schweizer, Robert Clark, and Claudia Vincenz at the University of Michigan.

\textbf{REFERENCES}


O’Farrell, P. H. (1975) \textit{J. Biol. Chem.} 250, 4007-4021


